DOCUMENT RESUME

ED 264 273 TM 850 759

AUTHOR Tatsuoka, Kikumi K.
TITLE A Latent Trait Model for Interpreting Misconceptions in Procedural Domains.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE Jul 82
CONTRACT N00014-79-C-0752
NOTE 19p.; In: Item Response Theory and Computerized Adaptive Testing Conference Proceedings (Wayzata, MN, July 27-30, 1982) (TM 850 744).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adaptive Testing; Elementary Secondary Education; *Error Patterns; *Evaluation Methods; Graphs; *Latent Trait Theory; *Mathematical Models; *Measurement Techniques; Scoring; Statistical Analysis; Tables (Data)
IDENTIFIERS *National Assessment of Educational Progress; Response Patterns

ABSTRACT

This study introduced a probabilistic model utilizing item response theory (IRT) for dealing with a variety of misconceptions. The model can be used for evaluating the transition behavior of error types, advancement of learning stages, or the stability and persistence of particular misconceptions. Moreover, it apparently can be used for relating the functional behavior of errors to other criterion measures such as creativity, anxiety, and motivation. One of several personal indices based on IRT was used to formulate the "rule space," which is a geometric representation of erroneous rules of operation. EC14, one of the indices, which is used primarily for detecting aberrant response patterns, appeared to be effective for separating clusters of response patterns from one another. A cluster of response patterns consists of the response pattern yielded by some rule and its "slips," due to partially consistent application of the rule. Using pattern classification to separate the clusters in the rule space accounts for variability of errors in the model. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

A Latent Trait Model for Interpreting Misconceptions in Procedural Domains

ED264273

Kikumi K. Tatsuoka
University of Illinois

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

The widespread availability of computers has brought a new dimension to the theory and practice of educational and psychological testing in the past several years. A new testing procedure that is not linear in form, adaptive testing, has been introduced and has demonstrated its superiority to conventional linear testing. Item response theory has proved to be a useful technique to measure the performance on such an individualized test, with scores obtained from the computerized adaptive testing procedure.

Another development remarkably different from the traditional psychometric method is that of error diagnostic testing, which can determine the sources of misconceptions as they occur along the steps necessary to reach the correct answer to a given problem. These approaches--fully utilizing the capability of computers--have successfully diagnosed hundreds of erroneous rules resulting from a variety of misconceptions, or incomplete knowledge, of one or two procedural steps in domains such as the arithmetic of whole numbers, signed numbers, and algebra (Brown & Burton, 1978; Matz, 1980; Shaw, Standiford, Klein, & Tatsuoka, 1982; Tatsuoka, Birenbaum, Tatsuoka, & Baillie, 1980; van Lehn, 1982).

It is a well-known fact among cognitive psychologists that correct answers can often be obtained for wrong reasons. Tatsuoka and her associates (Birenbaum & Tatsuoka, 1983; Tatsuoka & Tatsuoka, in press) have demonstrated that measuring students' performances by considering the total number of correct answers to be the score can be a serious mistake for achievement tests that are especially designed to measure the outcome of learning. According to Davis and McKnight (1980), Tatsuoka and Tatsuoka (1982), and van Lehn (1982), erroneous rules can produce the correct answer for a given item. For example, the correct answer to the problem $-16 - (-4)$ can be obtained by the following three erroneous rules: (1) subtracting the two numbers and taking the sign of the number with the larger absolute value; (2) changing the minus operation sign to addition, misunderstanding the parenthesis as the bars for absolute value and then applying the correct rule for addition; or (3) converting the subtraction to an addition problem by changing the sign of the second number, then subtracting the smaller absolute value from the larger absolute value and taking the sign of the first number for the answer. These three erroneous rules, which are committed by a substantial number of seventh graders, produce the correct answer (-12). Indeed, the third rule just described can produce the correct answers for all subtraction problems in which the first number has a larger absolute value than the second number.

TM 850 259

Obtaining the correct answers for the wrong reasons reduces the reliability of the test scores and leads to multidimensionality of data sets for achievement tests of the problem-solving type (Birenbaum & Tatsuoka, 1982; Tatsuoka & Birenbaum, 1981). The relationship between such psychometric properties and misconceptions has been investigated by Tatsuoka and her colleagues (Birenbaum & Tatsuoka, 1983; Tatsuoka & Tatsuoka, 1981, 1982a). This series of studies shows that when achievement tests are used as an integral part of instruction, performance on the tests indicate how well students master the use of the correct rule (taught by instruction) for responding to the items. If the data contain many responses resulting from consistent application of erroneous rules, then the dimensionality (in the factor analytic sense) of the data becomes larger. These studies imply that it is impossible to separate cognitive aspects of learning from psychometrics in order to interpret properly the students' performances on achievement tests when the tests are used to measure outcomes of instructional treatment.

Purpose

This study will introduce a measurement model using item response theory for dealing with the misconceptions committed by many students. It is useful to know the transitional behavior of error types which may be due to a change of instructional methods, to advancement of learning stages, or to the stability and persistance of particular misconceptions. Such knowlege helps to evaluate instruction, to measure the outcome of learning, and to obtain diagnostic information for designing remedial instruction depending on the types of misconceptions. The purpose of the model is to express such functional aspects of misconceptions quantitatively so that they can be related to other measures such as motivation or creativity.

Finally, a technique will be discussed for assessing erroneous rules when they are used inconsistently by a student due either to "slips" or to the instability of the students' misconceptions. The pattern classification technique to determine the student's latent state of knowledge or the underlying misconception seems very useful in handling the variability of errors.

Rule Space

Tatsuoka and Baillie (1982b) showed that all erroneous rules of operation in signed-number arithmetic were expressible as points in a geometric space called "rule space." In other words, rule space is a geometric representation of the rules used. Brown and Burton (1978), van Lehn (1982), Tatsuoka and Tatsuoka (1981), Birenbaum and Tatsuoka (1980), and Matz (1980) have found numbers of erroneous rules resulting from a variety of misconceptions in whole number subtraction problems, signed-number addition and subtraction problems, and algebra. Those rules are identified by closely examining responses obtained from a set of items that are carefully chosen so as to enable distinguishing each rule from others (Tatsuoka et al., 1980; van Lehn, 1982). For example, all erroneous rules of operation in signed-number addition and subtraction problems which have been discovered so far are represented by a unique set of responses to the set of items. Tatsuoka and Linn (1981) mapped a set of binary-score vectors of

these responses into the space spanned by the total scores and the values of the extended caution index, ECI4 (see below), along with the response vectors generated by 15 erroneous rules. In this space (referred to hereafter as the rule space), the erroneous rules resulting from the same kind of misconceptions clustered closely. This finding was confirmed in the domain of fraction addition problems (Tatsuoka & Chevalaz, 1983). Because of the nature of the extended caution index, unusual responses to the items are located in the lower or upper parts of the space and very usual responses scatter closely along the x-axis (the total scores or true scores).

Extended Caution Indices

A group of extended caution indices was introduced by Tatsuoka and Linn (1983), and their statistical properties were investigated in Tatsuoka and Tatsuoka (1982b). The values of the ECIs are calculated by first constructing two matrices; one is a binary score matrix

$$(y_{ij}),\ i = 1, \ldots, N,\ j = 1, \ldots, n \qquad [1]$$

where N is the number of students and n is the number of items in a test. The other is a probability matrix with elements P_{ij}, where P_{ij} is a logistic function of one, two, or three parameters.

In practice, the estimated P_{ij} obtained by substituting estimated item and person parameters in the logistic function can be used. One of the ECIs, ECI4, is defined as an index reflecting anomaly of an actual response pattern at a given level of ability θ_i. It is given by the complement of the ratio of two covariances: the numerator is the covariance of the ith row vector, y_i, in (y_{ij}) and the ith row, P_i, of the probability matrix (P_{ij}); the denominator is the covariance of the column-mean vector of $\underset{\sim}{G} = (G_{.1}, G_{.2}, \ldots, G_{.n})$ and the ith row vector P_i, both of (P_{ij}). Here

$$G_{.j} = \frac{1}{N} \sum_{i=1}^{N} P_{ij} \,. \qquad [2]$$

That is,

$$\text{ECI4} = 1 - \frac{\text{cov}(\underset{\sim}{y}_i\ ,\ \underset{\sim}{P}_i)}{\text{cov}(\underset{\sim}{G}\ ,\ \underset{\sim}{P}_i)} \,. \qquad [3]$$

The conditional expectation and variance of ECI4 are given by

$$E(\text{ECI4}|\theta_i) = 1 - \frac{\text{Var}(\underset{\sim}{P}_i)}{\text{cov}(\underset{\sim}{G}\ ,\ \underset{\sim}{P}_i)} \,, \qquad [4]$$

$$\mathrm{Var}(\mathrm{ECI4}|\theta_i) = \frac{\sum_{j=1}^{n} \sigma_{ij}^2 (P_{ij} - T_i)^2}{n^2 \mathrm{cov}^2(\underset{\sim}{G}, \underset{\sim}{P}_i)}. \qquad [5]$$

Thus, the standardized ECI4 is given by

$$\mathrm{ECI4}_z = \frac{n[\mathrm{cov}(\underset{\sim}{P}_i - \underset{\sim}{y}_i, \underset{\sim}{P}_i)]}{\sum_{j=1}^{n} \sigma_{ij}^2 (P_{ij} - T_i)^2} \qquad [6]$$

where

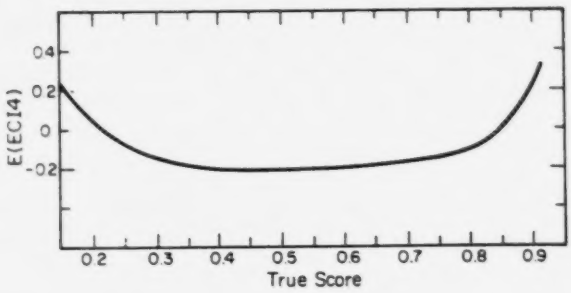
$$T_i = \frac{1}{n} \sum_{j=1}^{n} P_{ij} \qquad [7]$$

the raw-mean vector of (P_{ij}) and

$$\sigma_{ij}^2 = P_{ij}(1 - P_{ij}), \qquad [8]$$

the variance of item j at level i.

Figure 1
Expectation of ECI4 Against True Score



The figures of the expectations obtained from two different data sets, one being from a multiple-choice test and the other from a free-response item test, show an interesting difference. The expectation of the NAEP data set (a 68-item mathematics test used in the National Assessment of Educational Progress) shows a U-shaped curve (Figure 1), while that of the free-response test has a monotonically increasing curve (Figure 2). Figure 3 is the curve of the conditional standard error of ECI4 obtained from the same NAEP data set of 2,400 subjects. Tatsuoka and Tatsuoka (1982b) showed empirically that the standardized ECI4 has a normal distribution. It is not surprising because ECI4 is a weighted arithmetic mean of P_{ij}, $j = 1, 2, \ldots, n$.

Figure 2
Expectation of ECI4 Plotted Against the True Score

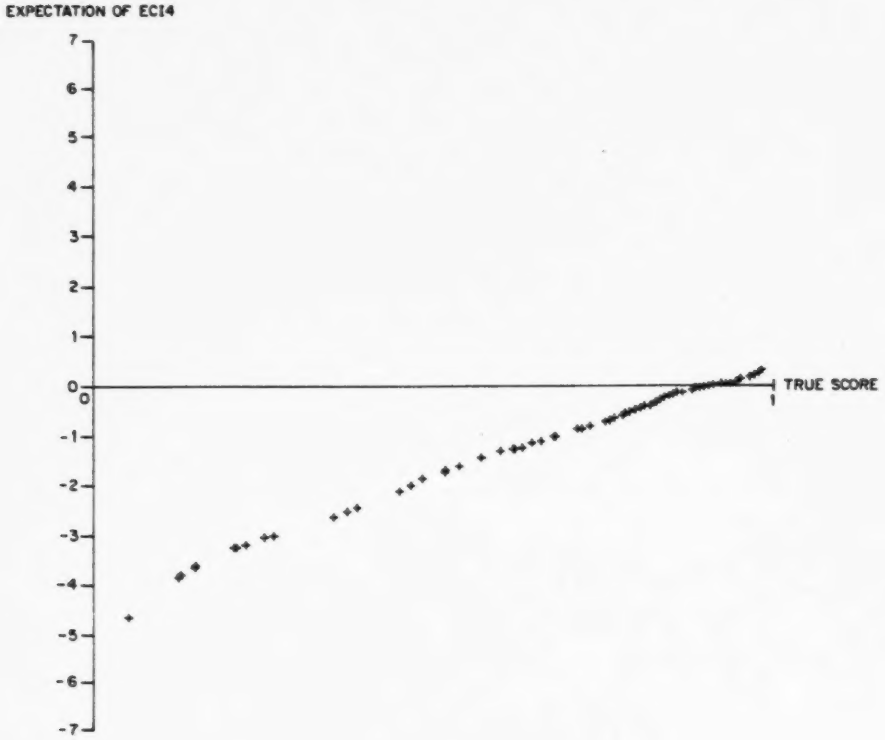
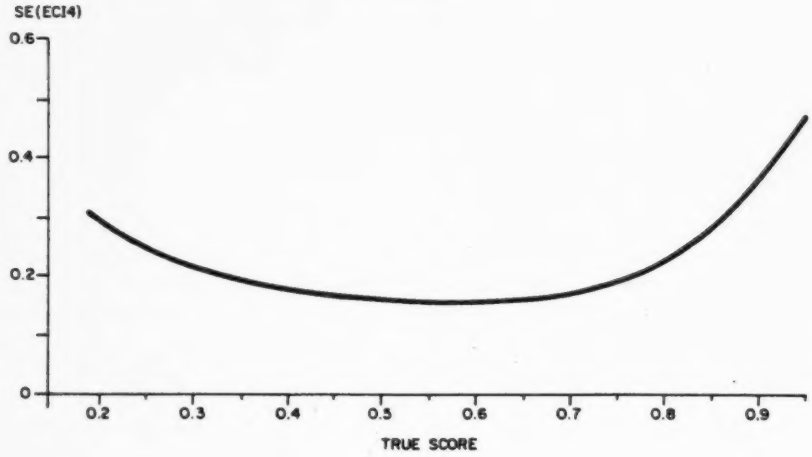


Figure 3
Standard Error of ECI4 for NAEP Data



A Map of Regions Classified by Misconceptions

One of the desirable properties of a rule space is that the responses resulting from the same kind of misconceptions are mapped into the neighborhood of the same kinds (Tatsuoka & Chevalaz, 1983; Tatsuoka & Linn, 1983). Tatsuoka and Chevalaz (1983) generated simulation data by controlling various sources of misconceptions in fraction addition problems and mapping them into the rule space. At the same time the responses generated by 52 erroneous rules described in Shaw, et al. (1982) were mapped into the same space along with the simulated responses. Figure 4 shows a map of those misconceptions.

Figure 4
Plot of True Score for Fraction Addition Test Against
Standardized Extended Index $ECI4_z$

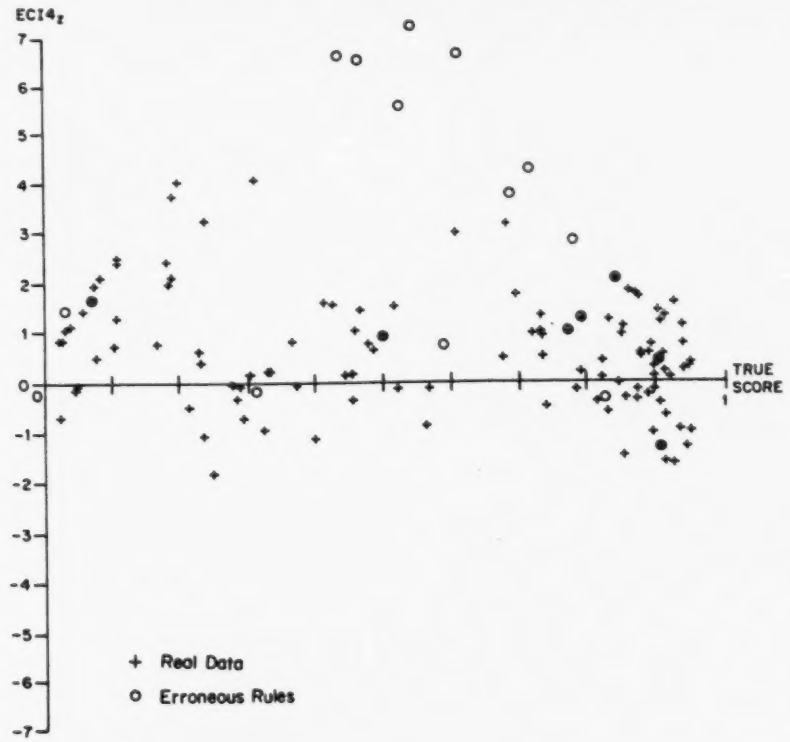


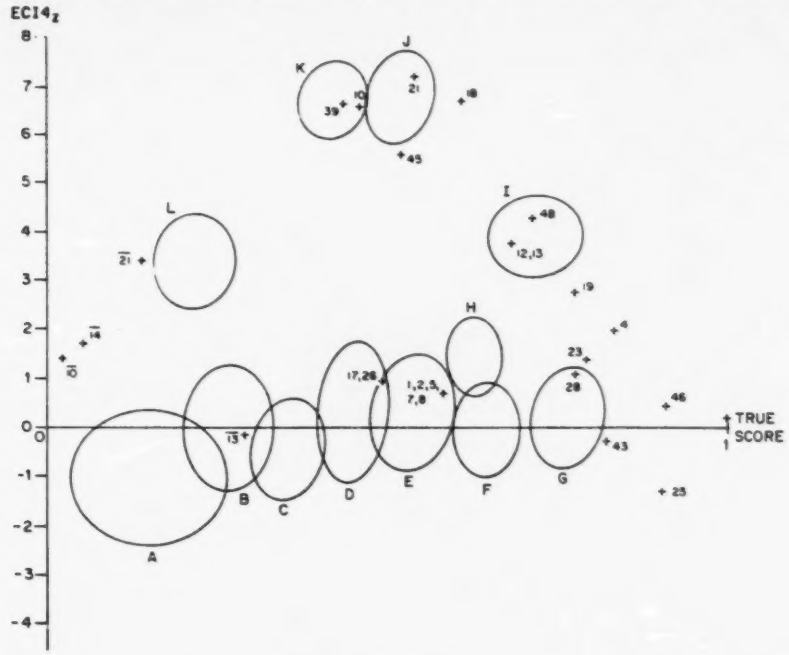
Table 1 describes the regions of A, B, ..., L in detail. For example, the response patterns (scored by the traditional scoring procedure) of Rules 39 and 10 fall in Region K, which is in the neighborhood of the responses consisting of correct answers for all mixed fraction problems and incorrect answers for others (or fraction-plus-fraction type). Rules 39 and 10 follows:

Rule 39: If the item was F+F type, then the numerator and denominator of the first fraction were added to obtain the numerator and the same was done for the second fraction to obtain the denominator. If the item was a mixed type, then the correct procedure was carried out.

Rule 10: A "1" was appended to all fractions (not mixed type), and then the resulting fraction was converted to an improper fraction. The correct procedure was then applied.

Rules 39 and 10 produce the correct answers for mixed types but always yield incorrect answers for F+F types. These responses result from some unusual erroneous rules; thus, the corresponding points are located at the upper middle section of the space in Figure 5.

Figure 5
Regions Determined by Various Misconceptions in the Fraction Addition Test and Erroneous Rules Described in Shaw et al. (1982)



On the other hand, the responses yielded by Rule 1 are correct answers for F+F types but incorrect for all mixed type items. Rule 1 is as follows:

Rule 1: The whole number was multiplied by the denominator and then the numerator was added. The denominator was not written, and the result was a whole number.

Figure 5 shows that the point in the rule space corresponding to Rule 1 belongs to Region E, whose points were generated as neighboring points of the responses of $x_j = 1$ if item j is an F+F type and $x_j = 0$ if otherwise.

It is interesting to note that the space in Figure 5 can be divided into several regions, each representing a unique source of misconception. In order to compare the regions in Figure 5 with actual students' responses from the fraction test, 186 students' response patterns are mapped into the space spanned by true scores and ECI4. Figure 4 shows their plots. By superimposing Figure 4 onto Figure 5, a student's underlying misconception, if there is one, can be

Table 1
Description of the 12 Regions Appearing in Figure 5

Region	No. of Items	Description of Region
A	8	If item j is of F+F type with equal denominators or if one denominator is divisible by the other, and the numerator of the answer is smaller than the denominator, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
B	12	If j is of F+F type and the denominators are equal, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
C	12	If j is of F+F type and the numerator of the answer is smaller than the denominator, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
D	22	If the denominators are equal, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
E	22	If j is of F+F type then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
F	20	If the answer needs to be converted, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
G	32	If j is of F+F type or has equal denominators, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
H	34	If j is of F+F type or has different denominators, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
I	10	If j is of mixed type and has equal denominators, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
J	18	If j has different denominators, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
K	10	If j is of mixed type, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.
L	10	If j is of F+F type and has different denominators, then $x_{ij} = 1$; otherwise, $x_{ij} = 0$.

diagnosed. If the student's response falls in Region B, then the description given in Table 1 provides the information about how the student responded to the items. However, not all responses fall in one of the regions without any ambiguity. Some responses may belong to a gray area between two regions, and it is difficult to describe specifically which region of the misconceptions is really the student's source of errors. It is necessary to apply the pattern classification technique and to establish the boundary lines between two neighboring regions.

The numbered points in Figures 4 and 5 corresponding to erroneous rules of operation in fraction addition problems (see Shaw, et al., 1982) are often

marked by two or more different numbers together. For example, Rules 1, 2, 5, 7, and 8 are represented by a single point. Rules 17 and 26 also share a single point. Since the true scores (x-axis) and the values of $ECI4_z$ (y-axis) are calculated from the binary response patterns obtained by the traditional "correct or incorrect" scoring procedure, these rules happened to produce the identical binary response patterns. In order to separate two different erroneous rules into different points in a geometric space, Tatsuoka and Baillie (1982) have proposed a new scoring procedure.

Component Response Patterns: Decomposing the Regular Scoring Procedure into Finer Components

Tatsuoka and Baillie (1982) demonstrated that several erroneous rules whose response patterns are identical with the traditional scoring procedure can be distinguished by decomposing the unit of the answer into finer components. Tatsuoka and Tatsuoka (1981) listed the response patterns of 45 erroneous rules in signed-number arithmetic also obtained from the regular scoring procedure. Some of the 45 binary patterns of 16 items were identical, although the descriptions of the erroneous rules which produced these identical patterns were not. There is no way to distinguish two such different rules just by examination of their binary response patterns.

However, if the regular scoring procedure is decomposed into finer components (e.g., the sign part of the answer for a given item and the absolute value part of the answer in signed-number problems or the whole number, numerator, and denominator parts of the answer for a fraction problem), then all the erroneous rules that have been discovered so far of the problem types in both signed-number and fraction arithmetic can be expressed uniquely as sets of the binary response patterns obtained from the component scoring procedure. For example, the response patterns of Rules 16, 32, 12, and 46 are given by:

Rule 16: The student subtracts the smaller absolute values from the larger absolute value and takes the sign of the number with the larger absolute value in his/her answer.

Rule 32: The two numbers are always subtracted as seen in Rule 16 but the "+" sign is always taken in the answers.

Rule 12: The student converts -2 - 8 and -13 - 5 into -2 + 8 and -13 + 5, respectively, but the other eight items are converted to addition problems correctly. Then the right addition rule is used to answer them.

Rule 46: The student has a strange idea about the parentheses and converts operation sign "-" to "+" first. Then he/she follows the rule: if the sign of the two numbers are minus, then change the sign of the second number to a "+"; if the signs of the two numbers are not alike, then the sign of the second number becomes a minus.

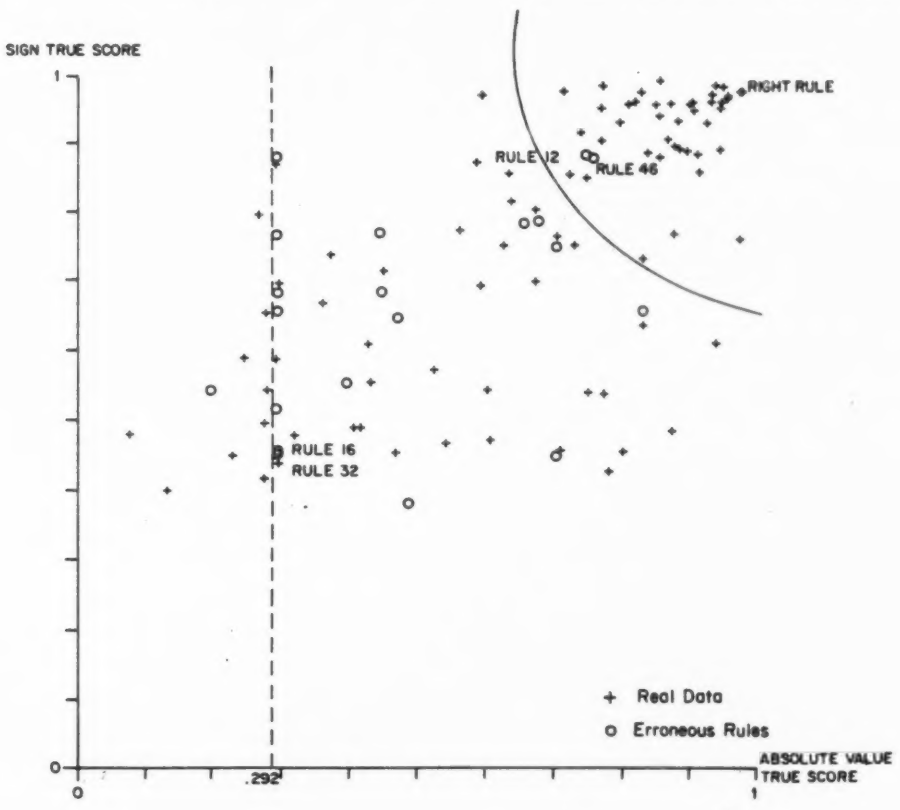
Table 2 shows that the regular response patterns for these four rules are elementwise products of the two component response patterns.

This new scoring method, the component scoring procedure, has been used in this study. Although the rationale of components scoring is based on a signed-number study, it can be generalized to other testing domains. For example,

Table 2
The Binary Response Patterns Generated by Four Rules and Three Different Scorings (R = Regular Scores, S = Sign-Component Scores, and A = Absolute-Value Component Scores)

	Rule 16				Rule 32				Rule 12				Rule 46			
Items	Response	R	S	A	Response	R	S	A	Response	R	S	A	Response	R	S	A
-3 - (-7) +4	-4	0	0	1	+4	1	1	1	+4	1	1	1	+4	1	1	1
-2 -8 = -10	+6	0	0	0	+6	0	0	0	+6	0	0	0	-10	1	1	1
5 - (-12) +17	-7	0	0	0	+7	0	1	0	+17	1	1	1	-7	0	0	0
-11 - +8 -19	-3	0	1	0	+3	0	0	0	-19	1	1	1	-19	1	1	1
9 - 4 = +5	+5	1	1	1	+5	1	1	1	+5	1	1	1	+5	1	1	1
-15 - (-9) = -6	-6	1	1	1	+6	0	0	1	-6	1	1	1	-6	1	1	1
-13 - 5 = -18	-8	0	1	0	+8	0	0	0	-8	0	1	0	-18	1	1	1
8 - (-6) +14	+2	0	1	0	+2	0	1	0	+14	1	1	1	+2	0	1	0
-5 - +11 -16	+6	0	0	0	+6	0	0	0	-16	1	1	1	-16	1	1	1
1 - 10 = -9	+9	0	0	1	+9	0	0	1	-9	1	1	1	-9	1	1	1

Figure 6
Plot of True Score for Absolute Value Component Against Sign Component



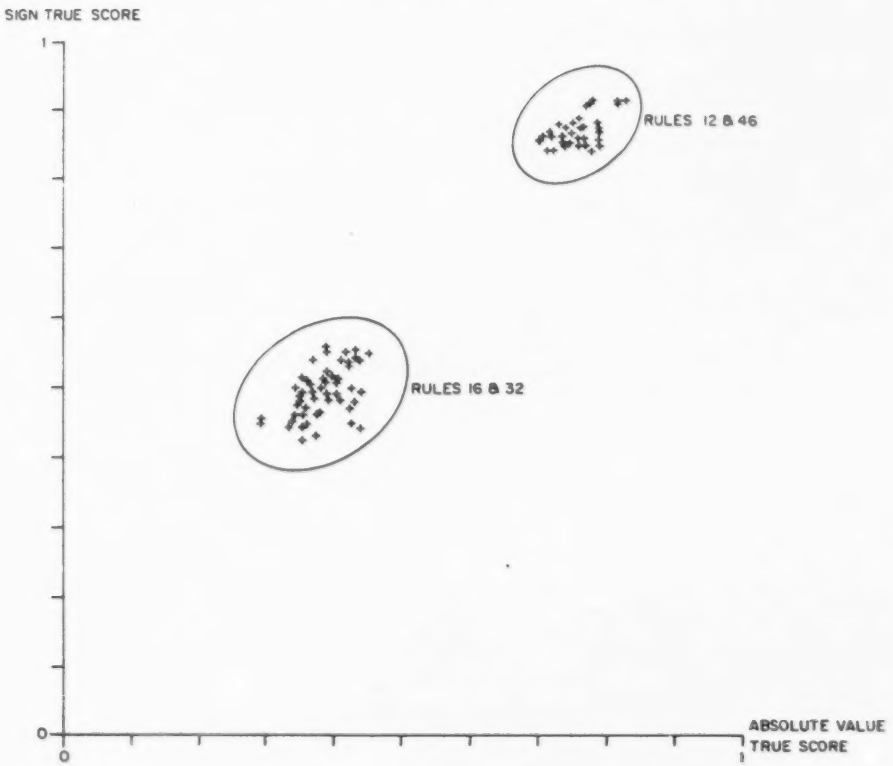
fraction problems can be decomposed into a whole number, numerator, and denominator parts.

Various Rule Spaces

A 40-item free-response test comprised of four parallel subtests of signed-number subtraction problems was administered to 172 eighth graders at a local junior high school. The traditional scoring of correct or incorrect answers was decomposed into a two-component scoring procedure of the absolute value and sign parts of the responses. The signs of the responses to the 40 items were scored correct or incorrect and so were the absolute values. The two-component response patterns were separately subjected to the estimation of item and person parameters of the 2-parameter logistic model. As was mentioned earlier, a rule space is defined as a geometric representation of the rules (including the correct rule and inconsistent application of two or more rules) used by the students.

The 20 small circles in Figure 6 represent 20 different erroneous rules, while the plus signs stand for a set of real responses to the 40 items for a

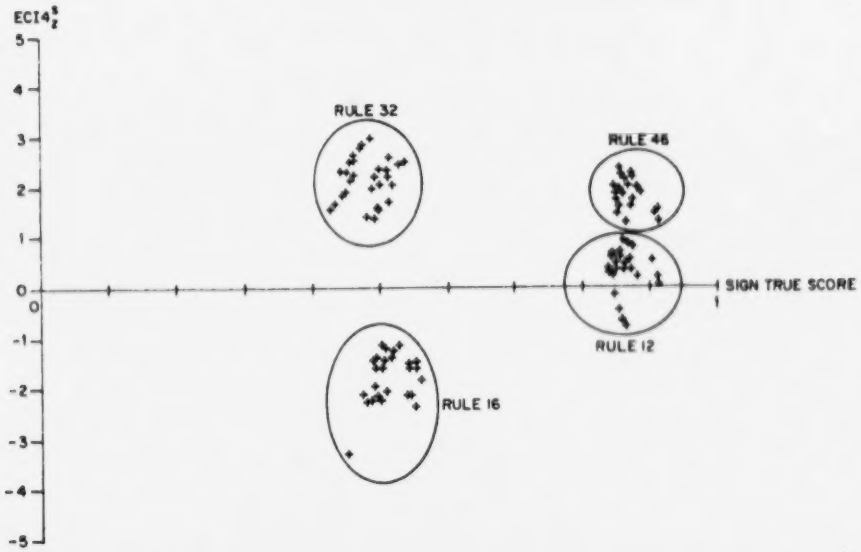
Figure 7
Plot of the Sign True Score Against Absolute Value for the Four Clusters Around Rules 12, 16, 32, and 46 in Figure 5



student. If the student responded to the 40 items by applying his/her erroneous rule consistently, then the student's point should coincide with the circle representing the rule. Figure 6 shows two such points. Most points do not show overlap; however, some real responses are located in the vicinity of a rule.

Tatsuoka and Baillie (1982) generated simulation data of responses resulting from inconsistent application of a rule. One or two items out of the 40 items do not follow a given erroneous rule and thus the component response patterns do not match completely with the patterns produced by the erroneous rule. Based on the 20 different rules, 20 sets of simulation data were generated and plotted on the space spanned by both the component true scores. As can be seen in Figure 7, Rules 16, 32, 12, and 46 and their simulated responses are not separated well because Rules 16 and 32, and 12 and 46, respectively, are already represented by very close circles in Figure 6. When plotted in terms of the sign true score against the standardized ECI4 obtained from the sign-component scores, four distinctly different clusters are formed in the new rule

Figure 8
Plot of the Clusters Around Rules 12, 16, 32, and 46 in Figure 5,
Sign True Score Against $ECI4^{S}_{Z}$



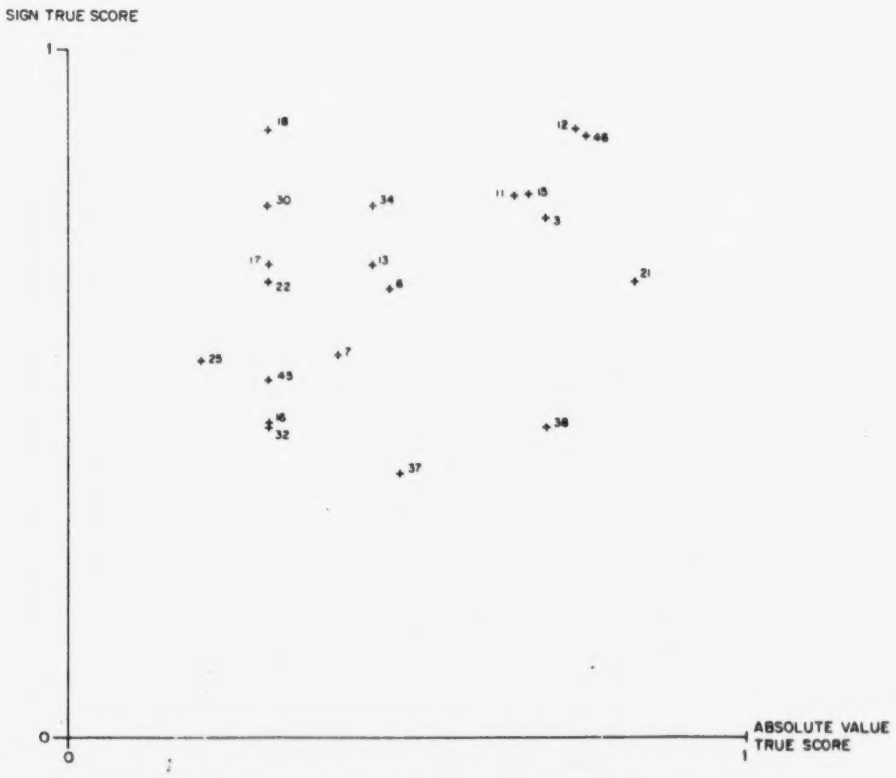
space, as shown in Figure 8. It is apparent that the values of ECIs are capable of separating response patterns that have either very close true scores or the same total scores.

Pattern Classification

In the previous section, Figure 8 showed the four erroneous rules; the nonconsistent responses neighboring each of them form four distinctly different clusters. By calculating a set of linear classification functions of the four clusters and by setting the boundaries to divide the four regions, it is possi-

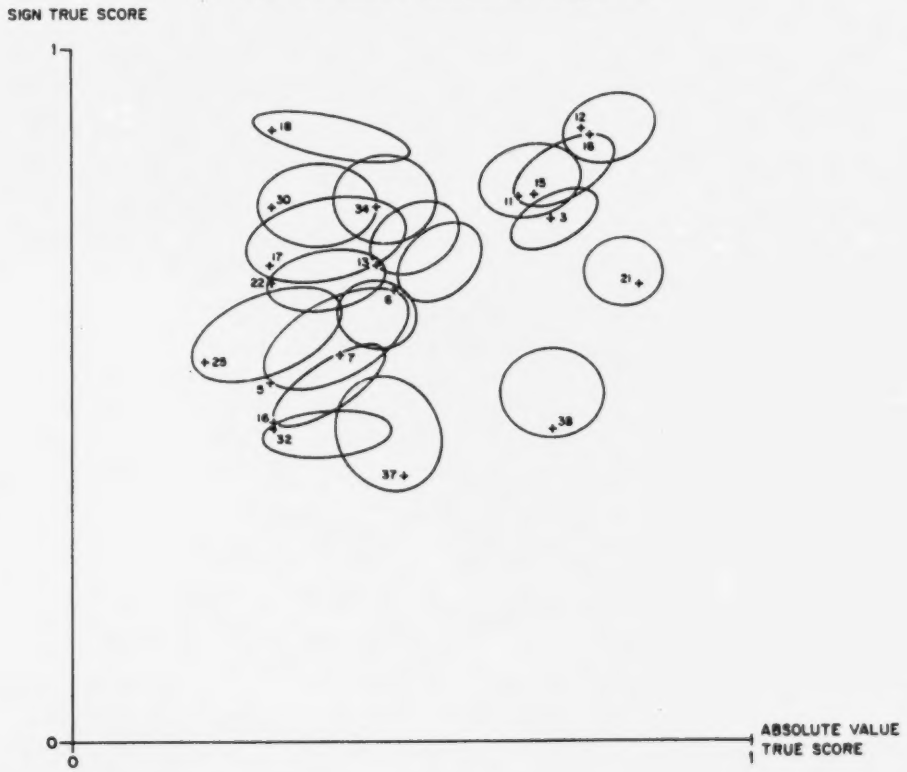
ble to identify the underlying misconception of a new response with some probability of error by examining in which region the new response falls. This is a procedure for pattern classification and recognition problems (Fukunaga, 1972). Thus, the problem of diagnosing an individual student's misconceptions has finally been transmuted into a classification problem. There are, however, many more erroneous rules than just four to which the classification problem can be applied, even in a simple arithmetic domain. Brown and Burton (1978) have discovered approximately 200 misconceptions, Tatusoka, et al. (1980), have found at least 30 erroneous rules in signed-number addition problems, and there are more erroneous rules in signed-number subtraction problems. A computer program called SIGNBUG has been developed for diagnosing these hundreds of erroneous rules in signed-number arithmetic tests (Tatsuoka & Baillie, 1982a). The logic of the mechanism is deterministic, however, and if a student responds to an item without using his/her rule, then SIGNBUG cannot determine the rule. A revised version of SIGNBUG can now diagnose a partial application of the student's rule to the test items, as when the student uses the correct rule for a subset of the test items and uses some erroneous rule for the rest of the items; but it is difficult to cope with responses yielded by random "slips." With the probabilistic approach of rule space and pattern classification, it is possible to remedy the weakness of the deterministic approach without losing the strength provided by SIGNBUG.

Figure 9
Twenty Erroneous Rules of Signed-Number Subtraction Problems



In general, pattern classification techniques must be applied to distinguish the 20 or more erroneous rules from one another. As has been mentioned earlier regarding fraction addition problems, the responses resulting from the same (or similar) kinds of conceptual mistakes tend to cluster closely. Moreover, the signed-number subtraction test demonstrated (Tatsuoka & Linn, 1983; Tatsuoka & Tatsuoka, 1981) a result very similar to the phenomenon seen in Figure 5. As a result, the pattern classification technique may be suitable for determining the source of misconceptions instead of considering a finer object--an erroneous rule of operation and its neighboring nonconsistent responses.

Figure 10
Twenty Clusters of the Responses Neighboring
the 20 Rules Shown in Figure 9



In this study, the focus was on determining the classification boundaries of 20 clusters neighboring each of the 20 erroneous rules of signed-number subtraction problems. The list of the 20 rules and their code numbers are given elsewhere (Tatsuoka & Tatsuoka, 1981) and are represented by the plots in Figure 9. Figure 10 shows the 20 clusters around each of the 20 rules. A stepwise discriminant analysis (BMDP7M) was used to determine the classification functions, and four independent variables--absolute value and sign true scores

Table 3

Classification Matrix: Number of Cases Classified into Group and Percent Correct

Rule	Percent Correct	Rule																			
		3	6	7	11	12	13	15	16	17	18	21	22	25	30	32	34	37	38	45	46
3	100	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	100	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	100	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	100	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	100	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	100	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	100	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0
16	100	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0
17	100	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0
18	100	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0
21	100	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0
22	100	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0
25	100	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0
30	100	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0
32	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0
34	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0
37	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0
38	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0
45	96.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0
46	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31
Total	99.8	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31

and $ECI4_z$ for absolute value and signs--were used in the analysis. The result is summarized in Table 3. Nineteen of the rules were perfectly classified without any error of classification, and only Rule 45 has one out of the 31 samples misclassified.

Summary and Discussion

This study introduced a probabilistic model utilizing item response theory for dealing with a variety of misconceptions. The model can be used for evaluating the transition behavior of error types, advancement of learning stages, or the stability and persistence of particular misconceptions. Moreover, it can be used for relating the functional behaviors of errors to other criterion measures such as creativity, anxiety, and motivation.

One of several personal indices based on item response theory was used to formulate the "rule space," which is a geometric representation of erroneous rules of operation. ECI4, one of the indices, which is used primarily for detecting aberrant response patterns, proved to be effective for separating clusters of response patterns from one another.

A cluster of response patterns consists of the response pattern yielded by some rule and its "slips," due to partially consistent application of the rule. Using pattern classification to separate the clusters in the rule space accounts for variability of errors in the model. It thus seems to be a promising approach toward assessing an individual's state of knowledge.

REFERENCES

Birenbaum, M., & Tatsuoka, K. K. The use of information from wrong responses in measuring students' achievement (Research Report 80-1-ONR). Urbana: University of Illinois, Computer-based Education Research Laboratory, 1980. (NTIS No. AD A097715)

Birenbaum, M., & Tatsuoka, K. K. On the dimensionality of achievement test data. Journal of Educational Measurement, 1982, 19, 259-266.

Birenbaum, M., & Tatsuoka, K. K. The effect of a scoring system based on the algorithm underlying the students' response patterns on the dimensionality of achievement test data of the problem solving type. Journal of Educational Measurement, 1983, 20, 17-26.

Brown, J. S., & Burton, R. R. Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 1978, 2, 155-192.

Davis, R. B., & McKnight, C. The influence of semantic content on algorithmic behavior. The Journal of Mathematical Behavior, 1980, 3, 39-87.

Fukanaga, K. Introduction to statistical pattern recognition. New York: Academic Press, 1972.

Matz, M. Toward a computational theory of algebraic competence. The Journal of Mathematical Behavior, 1980, 93-166.

Shaw, D. J., Standiford, S. N., Klein, M. F., & Tatsuoka, K. K. Error analysis of fraction arithmetic: Selected case studies (Research Report 82-2-ONR). Urbana: University of Illinois, Computer-based Education Research Laboratory, February 1982.

Tatsuoka, K. K., & Baillie, R. SIGNBUG: An error diagnostic computer program for signed-number arithmetic on the PLATO® system, 1982. (a)

Tatsuoka, K., K., & Baillie, R. Rule space, the product space of two score components in signed-number subtraction: An approach to dealing with inconsistent use of erroneous rules (Research Report 82-3-ONR). Urbana: University of Illinois, Computer-based Education Research Laboratory, 1982. (b)

Tatsuoka, K. K., & Birenbaum, M. The effect of different instructional methods on achievement tests. Journal of Computer-based Instruction, 1981, 8, 1-8.

Tatsuoka, K. K., Birenbaum, M., Tatsuoka, M. M., & Baillie, R. A psychometric approach to error analysis on response patterns (Research Report 80-3-ONR). Urbana: University of Illinois, Computer-based Education Research Laboratory, 1980.

Tatsuoka, K. K., & Chevalaz, G. M. A map representation of misconceptions: An approach utilizing item response theory and classification functions (Research Report 83-4-ONR-NIE). Urbana: University of Illinois, Computer-based Education Research Laboratory, May 1983.

Tatsuoka, K. K., & Linn, R. L. Indices for detecting unusual response patterns: Links between two general approaches and potential applications (Research Report 81-5-ONR). Urbana: University of Illinois, Computer-based Education Laboratory, August 1981.

Tatsuoka, K. K., & Linn, R. L. Indices for detecting unusual response patterns: Links between two general approaches and potential applications. Applied Psychological Measurement, 1983, 7, 81-96.

Tatsuoka, K. K., & Tatsuoka, M. M. Spotting erroneous rules of operation by the individual consistency index (Research Report 81-4-ONR). Urbana: University of Illinois, Computer-based Education Research Laboratory, 1981.

Tatsuoka, K. K., & Tatsuoka, M. M. Detection of aberrant response patterns. Journal of Educational Statistics, 1982, 7, 215-231. (a)

Tatsuoka, K. K., & Tatsuoka, M. M. Standardized extended caution indices and comparison of their error detection rates (Research Report 82-4-ONR). Urbana: University of Illinois, Computer-based Education Research Laboratory, March 1982. (b)

Tatsuoka, K. K., & Tatsuoka, M. M. Spotting erroneous rules of operation by the individual consistency index. Journal of Educational Measurement, in press.

van Lehn, K. Empirical studies of procedural flaws, impasses, and repairs in procedural skills (Technical Report ONR-8). Palo Alto CA: XEROX, Palo Alto Research Center, March 1982.

ACKNOWLEDGMENT

This research was sponsored by the Personnel and Training Research Program, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-79-C-0752. The authors gratefully acknowledge the assistance of Maurice Tatsuoka, Gerard Chevalaz, and John Eddins for editing and careful proofreading, and of Louise Brodie for typing and Roy Lipschultz for graphics. Some of the analyses in this report were done on the PLATO® System, which was developed by the University of Illinois. (PLATO® is a service mark of Control Data Corporation.)